Supplementary materials and methods

Genotyping

In brief, the 252 bp DNA fragment containing rs2270363 was amplified by PCR. And the unincorporated dNTPs and primers were removed with Phosphatase (SAP, Takara, Cat. No: 2720) and extranuclease I (ExoI, Takara, Cat. No: 2650A). The PCR products were then used as templates for single base extension (to genotype rs2270363), with the use of different fluorescence labeled ddNTPs. ABI 3730XL sequencer (Applied Biosystems, USA) was used to read out the genotype of the sample according to the color of the incorporated ddNTP and the results were analyzed by GeneMapper software (Applied Biosystems, USA). We used Sanger sequencing to validate genotyping results. In addition, SNPs failed to be genotyped with SNaPshot method were also genotyped with Sanger sequencing. All primer sequences are listed in **Supplementary Table 11**.

Dual-luciferase reporter gene assays

HEK-293T, SH-SY5Y and SK-N-SH cells were plated onto 96-well plates at a density of 2×10^4, 6×10^4 and 6.5×10^4 cells/well, respectively. The pRL-TK (internal control, 50 ng, Promega, Cat. No: E2241) and the recombinant vectors (150 ng) were co-transfected into the cells using the Lipofectamine^{TM} 3000 transfection reagent (Invitrogen, Cat. No: L3000-015). 48 hours post-transfection, the luciferase activity was determined with the Dual-Luciferase Reporter Assay System (Promega, Cat. No: E1980) and the Luminoskan Ascent chemiluminesence analyzer (Thermo scientific, Thermo Luminoskan Ascent). All primer sequences are listed in **Supplementary Table 12**.

Electrophoretic mobility shift assay

The 3’ end of the oligonucleotides were labeled with biotin using the EMSA Probe Biotin Labeling Kit (Beyotime, Cat. No: GS008), and single-strand oligonucleotides were annealed to form double-stranded complementary sequences. The Chemiluminescent EMSA Kit (Beyotime, Cat. No: GS009) was used to perform the EMSA according the manufacturer’s instructions. Briefly, 50 fM probes were incubated with nuclear extracts for 20 minutes. The resultant incubation solution were then subjected to electrophoresis on 6.5% non-denaturing polyacrylamide gels with 0.5 × Tris borate/EDTA and electro-transfer to nylon membrane (Millipore, Cat. No: INYC00010). After UV cross-linking, the membrane was incubated with Streptavidin-HRP Conjugate and images were taken using the Mini chemi610 imaging system (SAGECREATION).

Knockdown and overexpression experiments

The short hairpin RNAs (shRNAs) were designed using two online shRNA design tools (https://rnaidesigner.thermofisher.com/rnaiexpress/sort.do or https://www.sigmaaldrich.com/china-mainland/zh/life-science/functional-genomics-and-rnai/shrna/individual-genes.html). shRNAs targeting *USF1* (human), *MAX* (human) and *Nmral1* (mouse) were cloned into pLKO.1-EGFP-Puro vector. And shRNAs targeting *Nmral1* (rat) were ligated to pSicoR-Ef1a-mCh-puro vector. HEK293T were co-transfected with 10 μg shRNA pLKO.1-EGFP-Puro vector or *NMRAL1*-OE-plenti-CAG -IRES-GFP, 5 μg psPAX-2 and 2 μg pMD2.G. Lentiviral particles were concentrated using Lenti-X Concentrator (Takara, Cat. No: 631231) for infecting mNSCs. 48 hours post-infection, the cells that stably expressing shRNA were selected with 2 μg/mL puromycin (Sigma, Cat. No: 540222) for 14 days. The cells that stably expressing *NMRAL1* were selected with 200 μg/mL G418 (Sangon biotech, Cat. No: B540723-0005) for 7 days.

RNA sequencing

Transcriptome of the wild type mNSCs and *Nmral1* knocked-down mNSCs were sequenced as previously described[1, 2]. Total RNA was isolated using TRIzol reagents (Life technologies, Cat. No: 15596018). RNA purity was assessed with the kaiaoK5500®Spectrophotometer (Kaiao, China). RNA integrity and concentration were evaluated using the RNA Nano 6000 Assay Kit of the Bioanalyzer 2100 system (Agilent Technologies, USA). RNA sequencing data were generated with the Illumina NovaSeq 6000 platform. DESeq2 was used to identify the differentially expressed genes (DEGs).[3] Gene Ontology (GO) enrichment analysis and Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment analysis were then performed using Metascape (https://metascape.org).[4]

5-Ethynyl-2'-deoxyuridine (EdU) proliferation and CCK-8 assays

For the EdU proliferation assay, 1.5×10^5 mNSCs cells were plated onto 24-well plates which were pre-coated with Laminin (Sigma, 20 μg/mL, Cat. No: L2020-1mg). After culturing for 48 hours, EdU (RIBOBIO, Cat. No: C00053) was added to a final concentration of 20 μg/mL and incubated for 1 hour. The EdU positive cells were detected using The Cell-Light EdU Apollo567 In Vitro Kit (RIBOBIO, Cat. No: C10310-1) strictly according to the instructions.

For CCK-8 assay, 1.8×10^4 mNSCs cells were seeded onto 96-well plates which were pre-coated with Laminin (Sigma, 20 μg/mL, Cat. No: L2020-1mg). After incubating for 0, 24, 48 and 72 hours, 10 μL CCK-8 (Beyotime, Cat. No: C0042) reagent was added to each well and incubated for 2 hours. The absorbance was then detected at 450 nm with a Spectrophotometer (BioTek, USA).

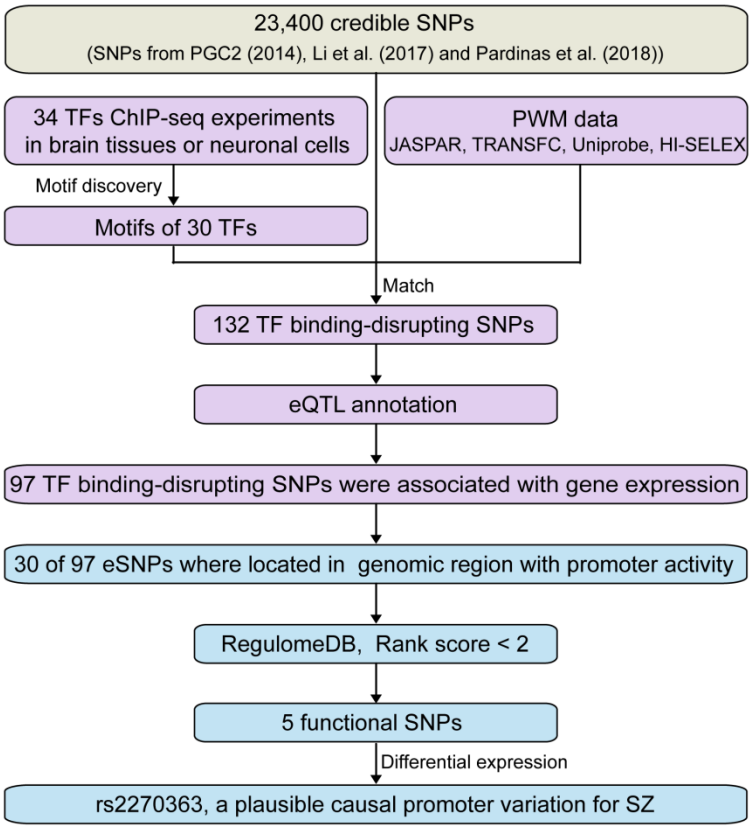
Immunofluorescence staining

The mNSCs cells were fixed with 4% Paraformaldehyde (PFA) for 15 min and

permeabilized using PBS with 0.3% Triton (PBST) for 15 min at room temperature. The cells were then blocked with blocking buffer (Beyotime, Cat. No: P0102) at room temperature for 1 hour. Subsequently, primary antibodies were incubated at 4 °C for overnight. Secondary antibodies were then incubated at room temperature for 1 hour (in dark room). Before taking images, the cells were counterstained with DAPI for 15 min, and photographed using FluoViewTM FV1000 Laser Scanning Confocal Microscopy (Olympus) and LSM 880 Basic Operation (Carl Zeiss, Germany). The primary antibodies used in this study were: PAX6 (Milipore, 1:500, Cat. No: ab2237), NESTIN (Chemicon, 1:500, Cat. No: mab353), SOX2 (Santa Cruz, 1:200, Cat. No: sc-17320), GFAP (Sigma, 1:1000, Cat. No: G9269), MAP2 (Millipore, 1:200, Cat. No: AB5622), Chicken polyclonal GFP antibody (Abcam, Cat. No: ab13970), and Rabbit polyclonal mCherry antibody (GeneTex, Cat. No: GTX128508).

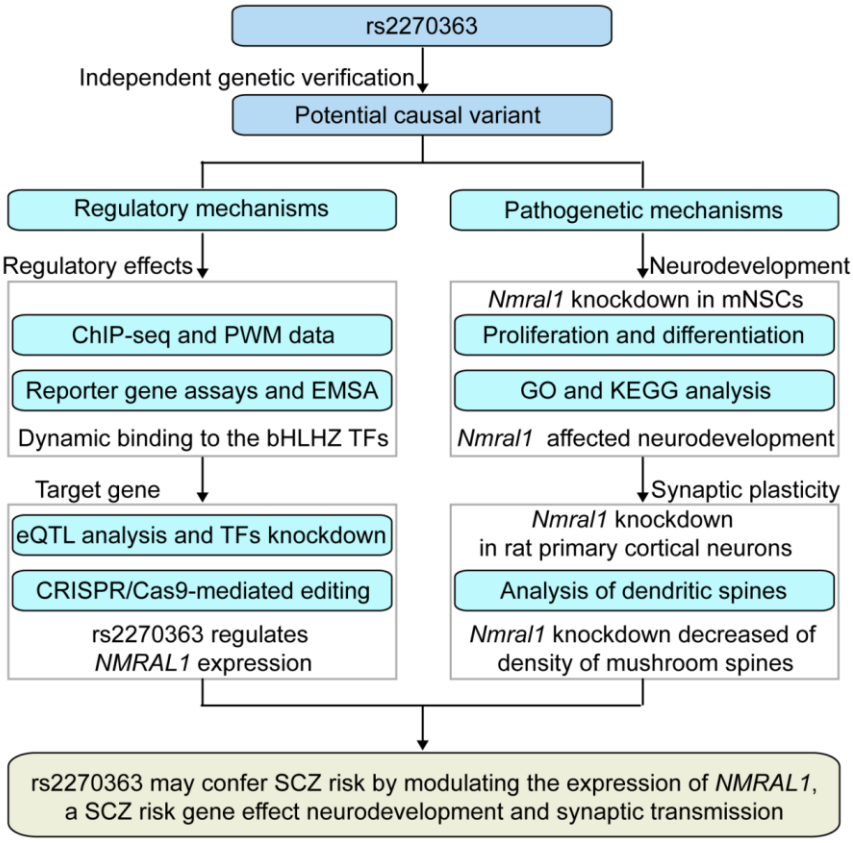
Meta-analysis

A meta-analysis was performed by combining our Chinese samples with the EUR and EAS samples.[5] A total of 60,136 cases and 86,647 controls were included in the meta-analysis, including the EUR samples (33,640 cases and 43,456 controls), the EAS samples (22,778 cases and 35,362 controls) and our Chinese samples (3,718 cases and 7,829 controls) (**Table 1**).

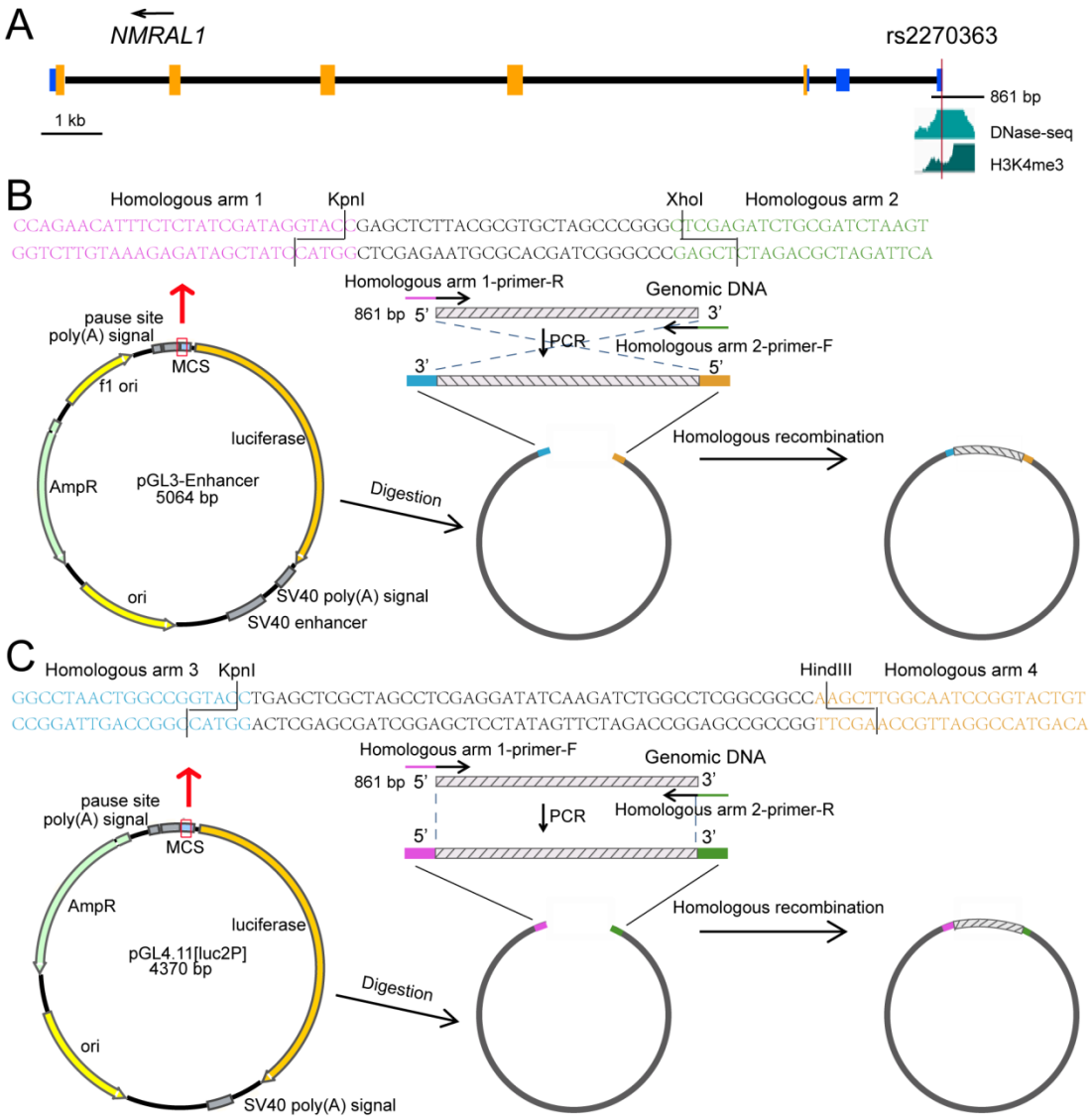


Supplementary Figure 1. Functional SNP screening prioritized rs2270363 as a potential causal variant. Our previous functional genomics study identified 132 TF binding-disrupting risk SNPs with functional consequences.[6] Among these SNPs, 97 showed associations with gene expression in the human brain (these SNPs are called eSNPs). And 30 out of 97 eSNPs are located in the promoter region and marked with H3K4me3 (a marker for active promoters) in UCSC genome browser (http://genome.ucsc.edu/), implying that these SNPs may affect promoter activity. We further used RegulomeDB and pinpointed 5 high-confidence (RegulomeDB rank score < 2) functional SNPs.[7] Finally, differential expression of the eQTL genes of the 5 SNPs (located in the promoter region) were examined in schizophrenia cases and controls.[8] The expression of the eQTL genes of rs223387 did not show significant change in SZ cases and controls. We noticed that rs2269524 is in high linkage disequilibrium (r^2 > 0.95) with rs1801311, a functional SNP whose regulatory mechanisms in schizophrenia have been recently reported.[1] Thus, we did not investigate rs1801311 in this study. Two eQTL genes (*FAM109B* and *LINC00634*) of rs2269524, two eQTL genes (*NDUFA2* and *SRA1*) of rs3822346, and two eQTL genes (*NMRAL1* and *CORO7*) of rs2270363 showed differential expression in SZ cases compared with controls (Supplementary Table 8). However, we noted that three SNPs (rs2269524, rs1801311 and rs3822346) are far away from their differentially expressed eQTL genes and are not located in the promoter regions of their differentially expressed eQTL genes, suggesting that they are unlikely regulate the expression of these genes by affecting the promoter activity. For example, our previous study showed that rs1801311, which is located in the first exon of *NDUFA6*,

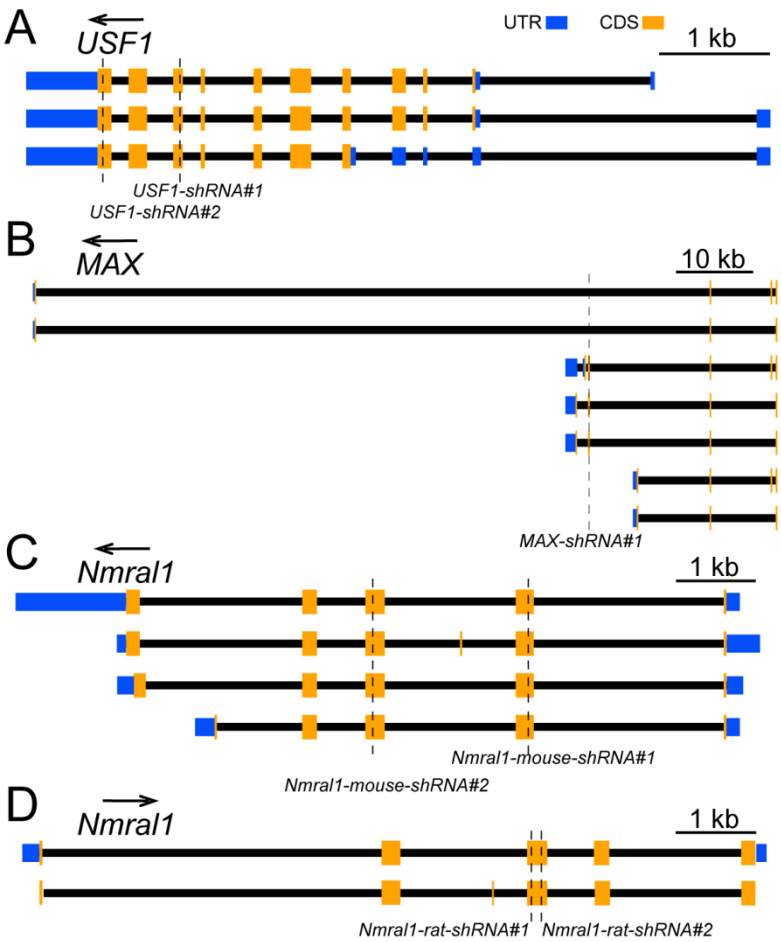
may confer schizophrenia risk by regulating the expression of the distal gene *NAGA* (an eQTL gene of rs1801311) rather than *NDUFA6*[1]. Interestingly, rs2270363, which is located in the *NMRAL1* promoter region, was associated with *NMRAL1* expression in the human brain, and *NMRAL1* expression was down-regulated in SZ cases compared with controls. This suggests that rs2270363 may regulate the expression of *NMRAL1* by modulating promoter activity. These lines of evidence suggested rs2270363 as a plausible causal variant for SZ at the 16p13.3 risk locus.



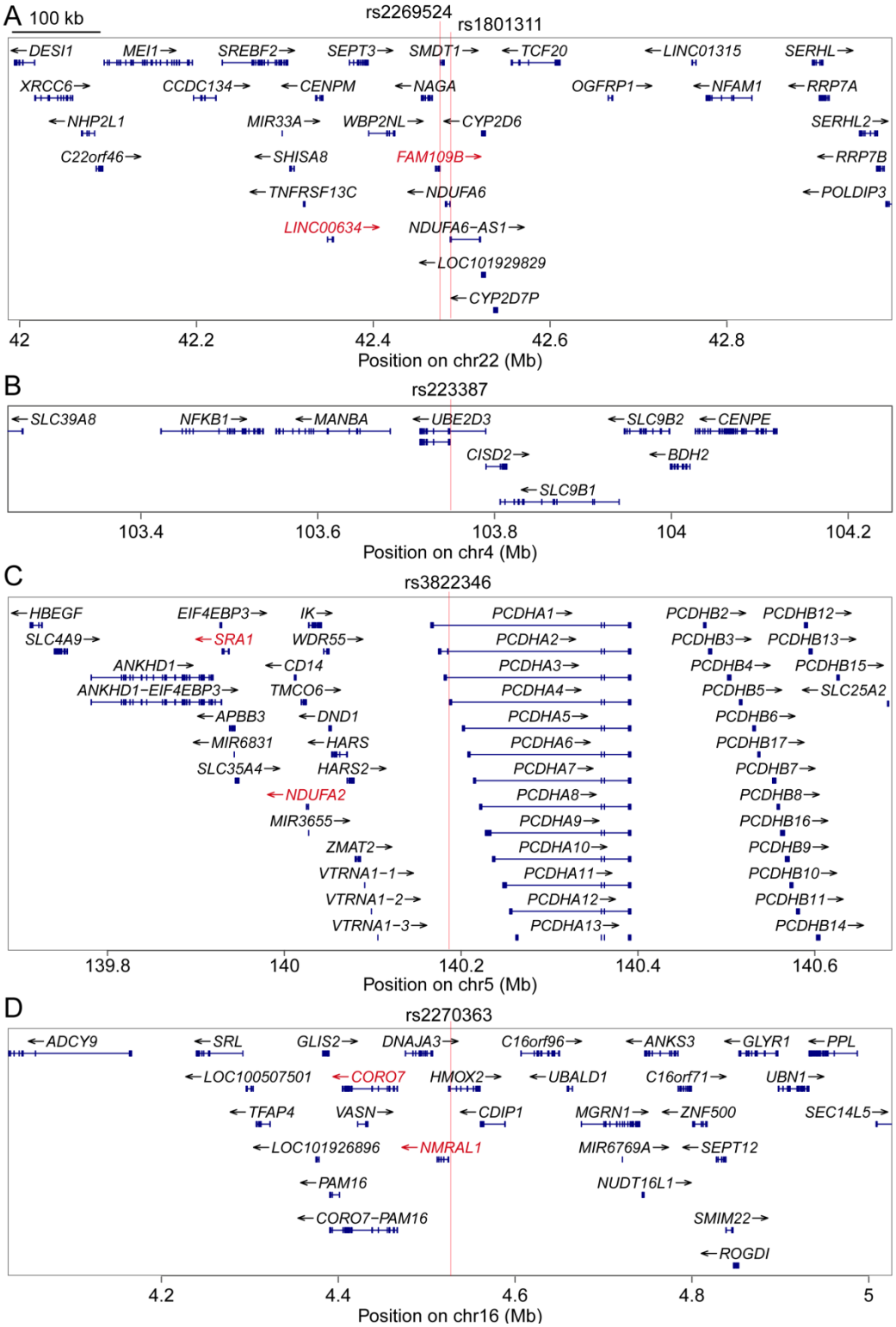
Supplementary Figure 2. Overview of this study. In this study, we firstly verified that rs2270363 was significantly associated with SCZ in China population. To explore the regulatory mechanisms of rs2270363, we leveraged a series of assays, including reporter gene assays and EMSA, as well as TF knockdown and CRISPR-Cas9-mediated genome editing. We showed that rs2270363 plays a role in regulating transcriptional activity through dynamic binding to the bHLHZ proteins. Besides, we identified the potential target gene of rs2270363 using eQTLs analysis, CRISPR/Cas9-mediated editing and TFs knockdown. Further functional studies showed that *Nmral1* knockdown affects proliferation and differentiation of neural stem cells, pathways related to neurodevelopment, and density of mushroom spines. In summary, our study demonstrates that rs2270363 may confer SCZ risk by modulating the expression of *NMRAL1*, a SCZ risk gene that affects neurodevelopment and synaptic transmission.



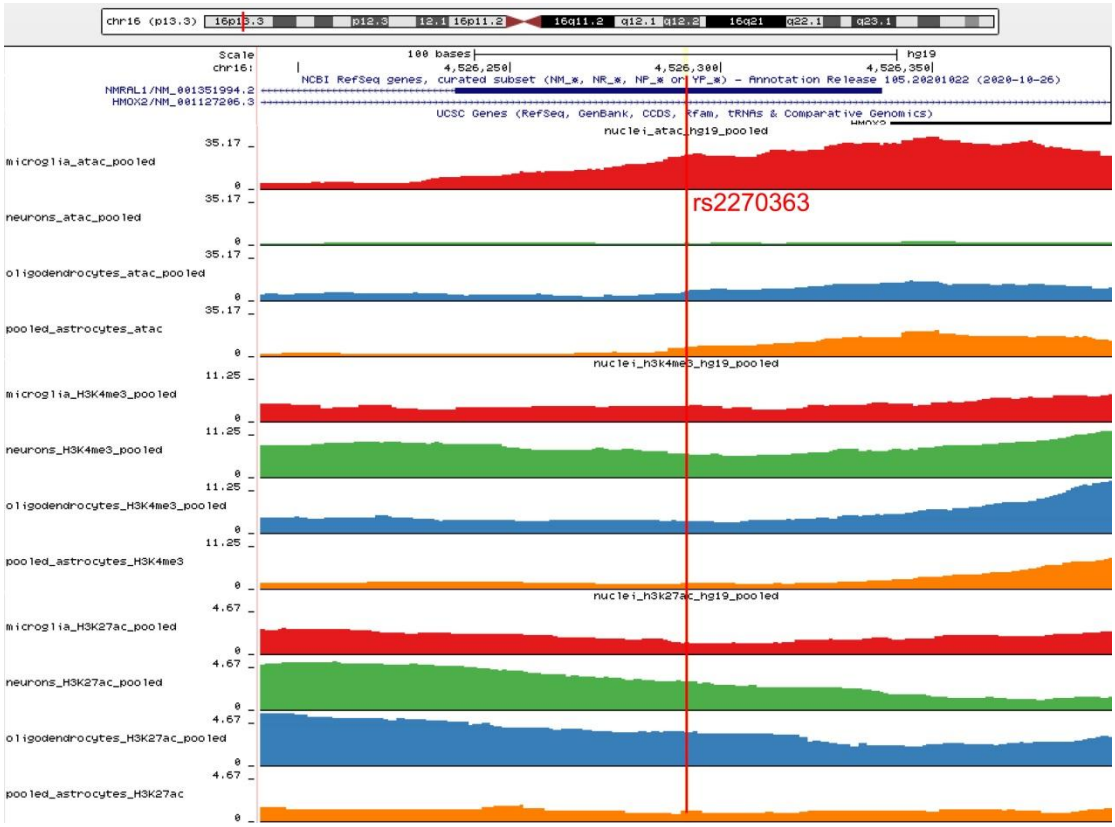
Supplementary Figure 3. Vectors design of dual-luciferase reporter gene assays. (**A**) The 861 bp DNA fragment (containing rs2270363), which was marked with H3K4me3 and located in open chromatin regions in SK-N-SH cell lines, was selected for dual-luciferase reporter gene assays. (**B**, **C**) Diagram of the construction of the recombinant vectors for reporter gene assays. The orientation of the fragment in pGL3-Enhancer was the same as that of the transcription of the *NMRAL1* for **A**. The orientation of the fragment in pGL4.11 was opposite to that of the transcription of the *NMRAL1* for **B**.



Supplementary Figure 4. The regions targeted by the designed shRNAs. (**A**, **B**) The regions targeted by the *USF1* and *MAX* shRNAs (human). (**C**) The regions targeted by the *Nmral1* shRNAs (mouse). (**D**) The regions targeted by the *Nmral1* shRNAs (rat).

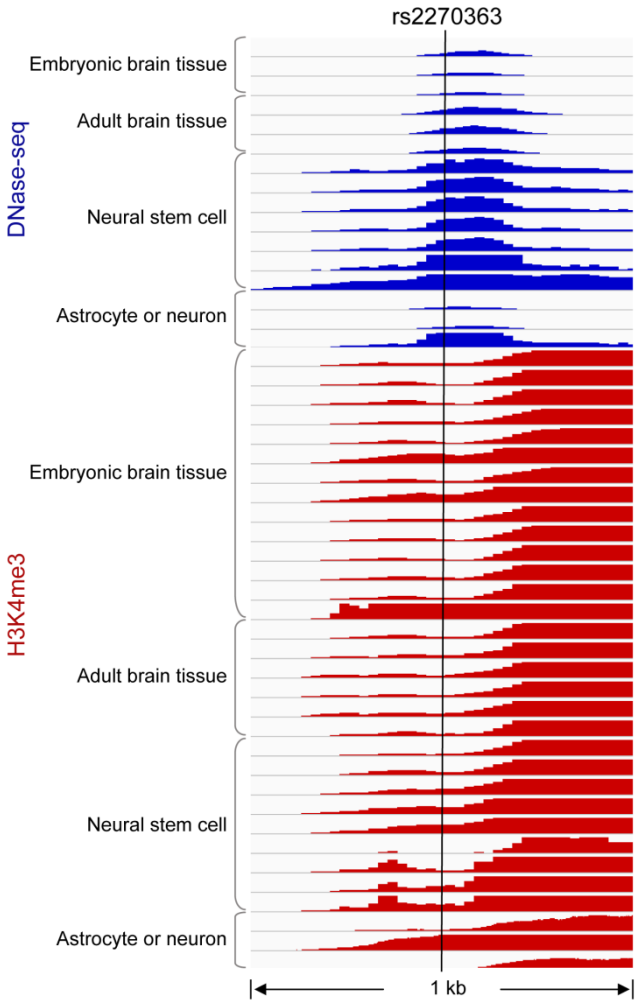


Supplementary Figure 5. The genomic region (1 Mb) containing 5 high-confidence functional SNPs. (**A**) rs1801311 is located in the *NDUFA6* promoter region, and rs2269524 is located in the *SMDT1* promoter region. (**B**) rs223387 is located in the *UBE2D3* promoter region. (**C**) rs3822346 is located in the *PCDHA4* promoter region. (**D**) rs2270363 is located in the *NMRAL1* promoter region. rs2269524, rs1801311 and rs3822346 are far away from their differentially expressed eQTL genes (red font) in the genome.

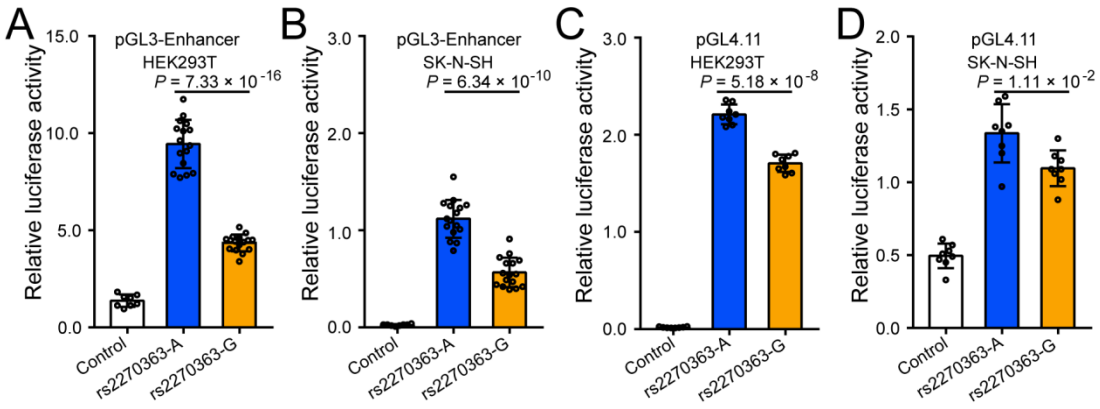


Supplementary Figure 6. rs2270363 is located in a genomic region marked with H3K4me3 (a marker for promoters) and H3K27ac (a marker for enhancer) signals in human microglia, neurons and oligodendrocytes. Data were from Nott et al.[9]

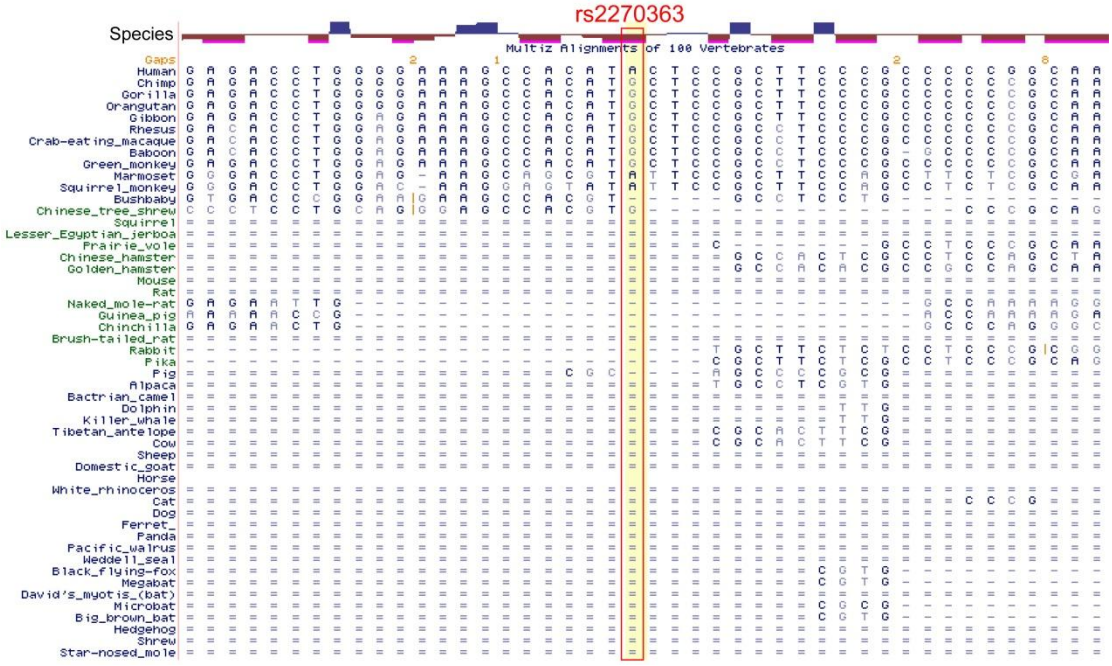
Supplementary Figure 7. rs2270363 is located in a genomic region marked with H3K4me3 (a marker for promoters) and H3K27ac (a marker for enhancer) signals in human brain tissues and neural cells (including neuroblastoma cell lines). Red indicates high H3K4me3 signal, yellow indicates H3K27ac signal, green indicates DNase signal, and blue indicate CTCF signal. Data were from ENCODE.[10]



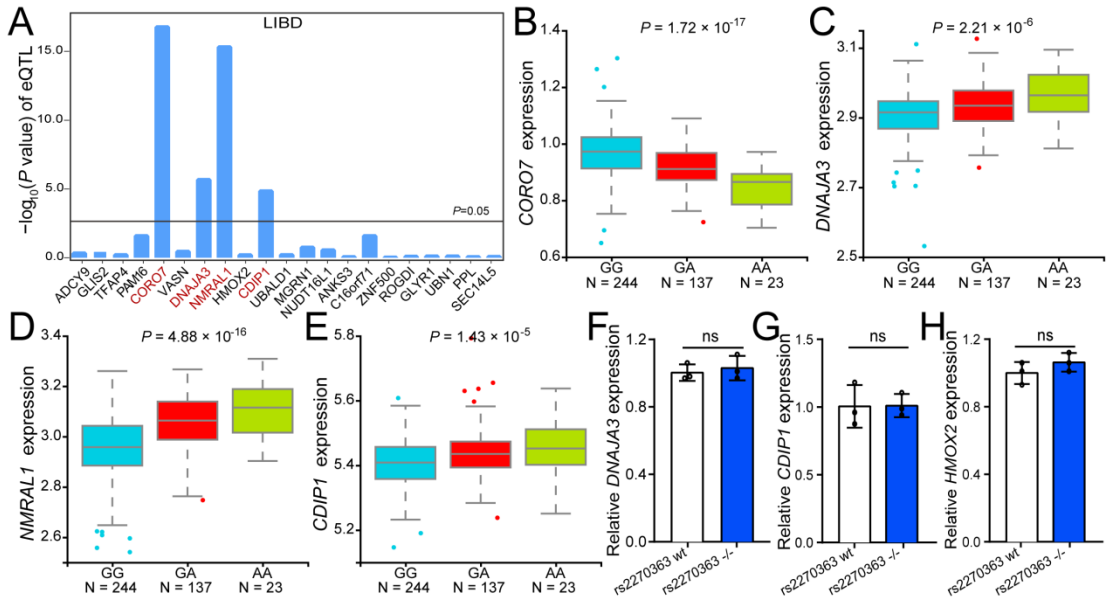
Supplementary Figure 8. DNase-Seq and H3K4me3 signals surrounding rs2270363 in human brain tissues and neural cells. Data were from ENCODE.[10]



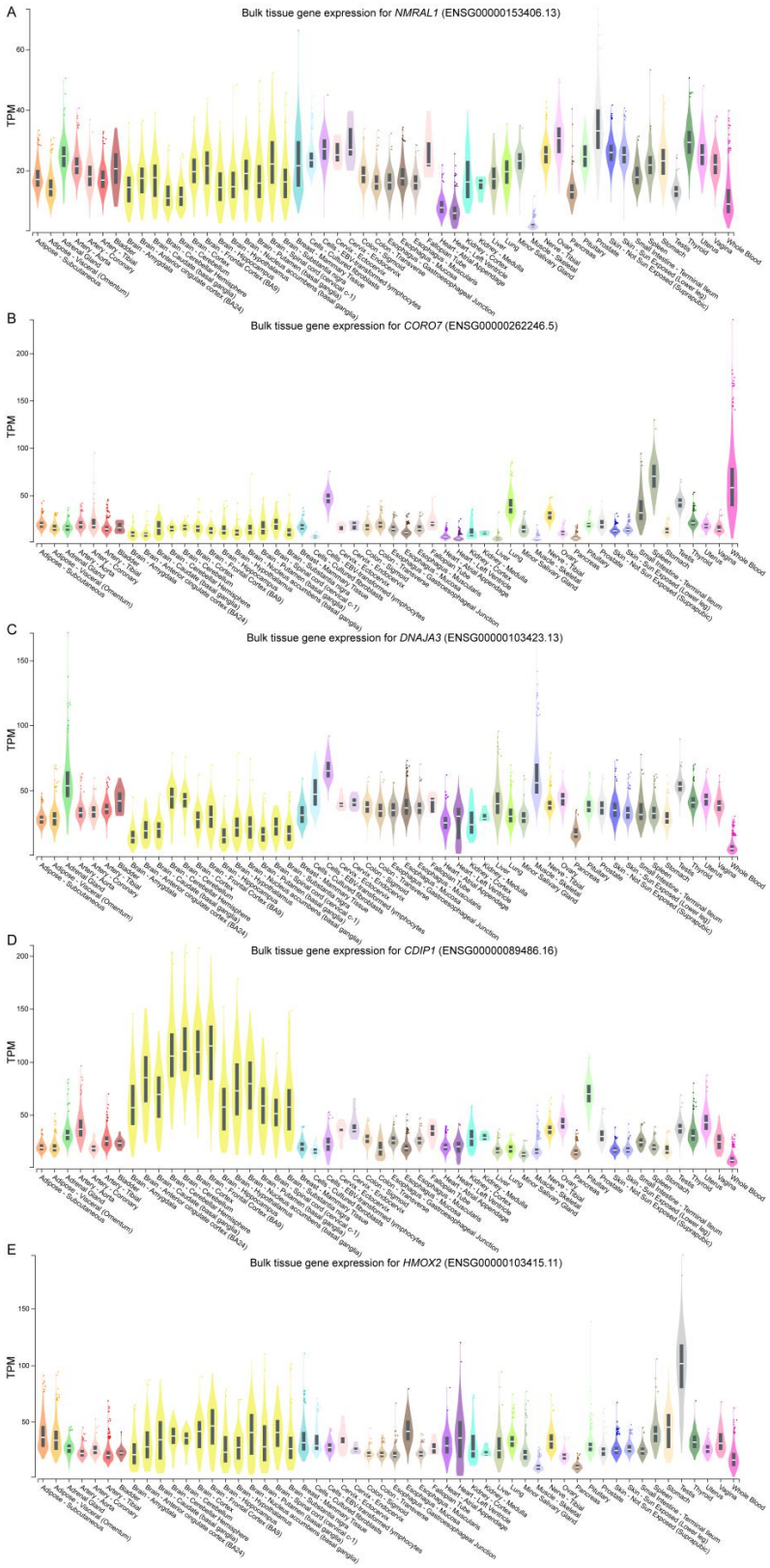
Supplementary Figure 9. Dual-luciferase reporter gene assays. (A-D) Dual-luciferase reporter gene assays for HEK293T and SK-N-SH cells. HEK293T and SK-N-SH cells were transfected with pGL3-Enhancer control vector and pGL3-Enhancer recombinant vector containing rs2270363 alleles for **A** and **B**. HEK293T and SK-N-SH cells were transfected with pGL4.11 control vector and pGL4.11 recombinant vector containing rs2270363 alleles for **C** and **D**. n = 8 for control group in **A** and **B**, n = 16 for experimental group in **A** and **B**, n = 8 for **C** and **D**. Unpaired two-tailed *Student's t*-test was used to test if the difference reached significance level ($P = 0.05$).



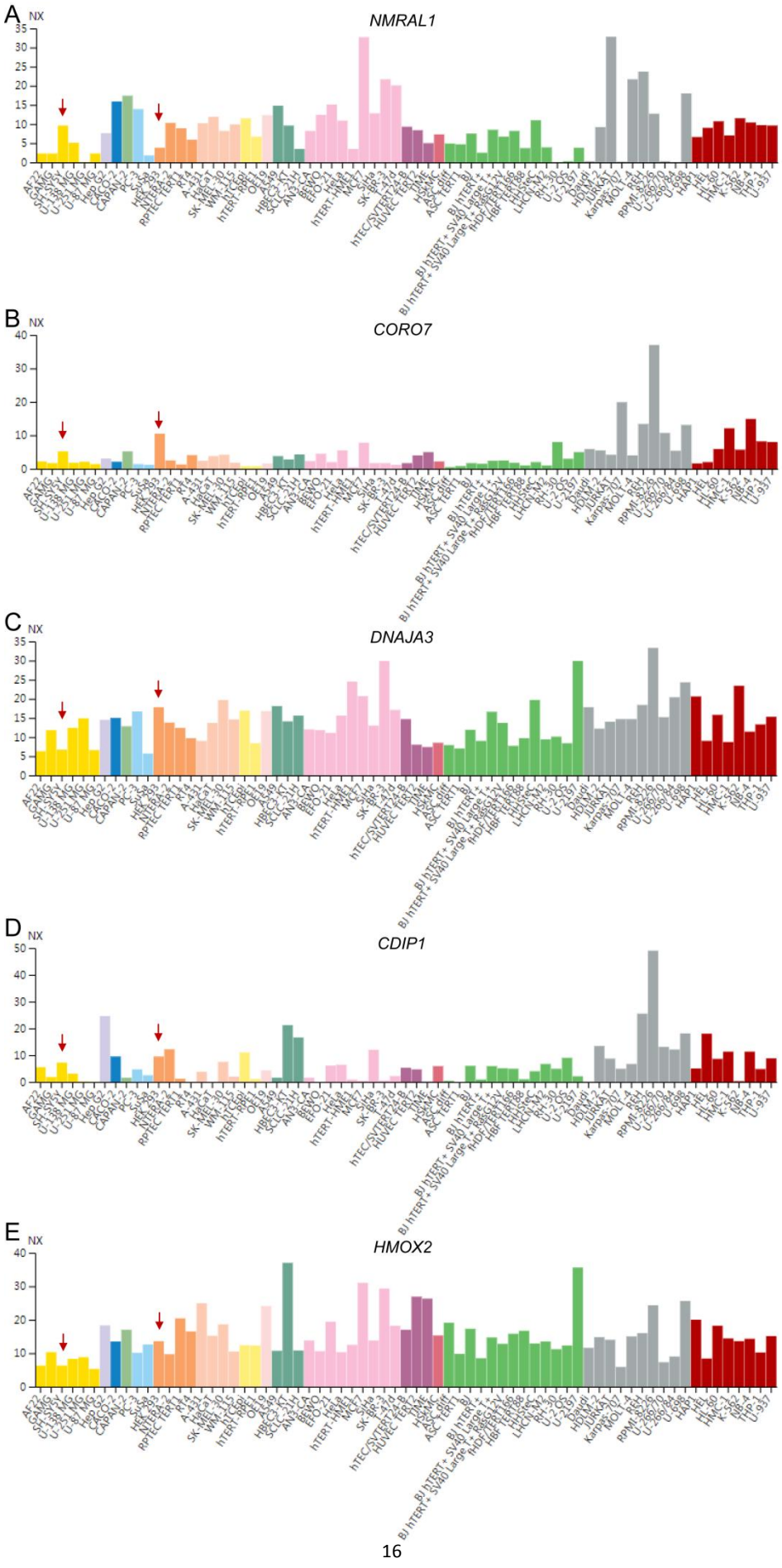
Supplementary Figure 10. Sequence conservation of genomic region containing rs2270363. rs2270363 was marked by red box. Data were from UCSC genome browser (http://genome.ucsc.edu/).



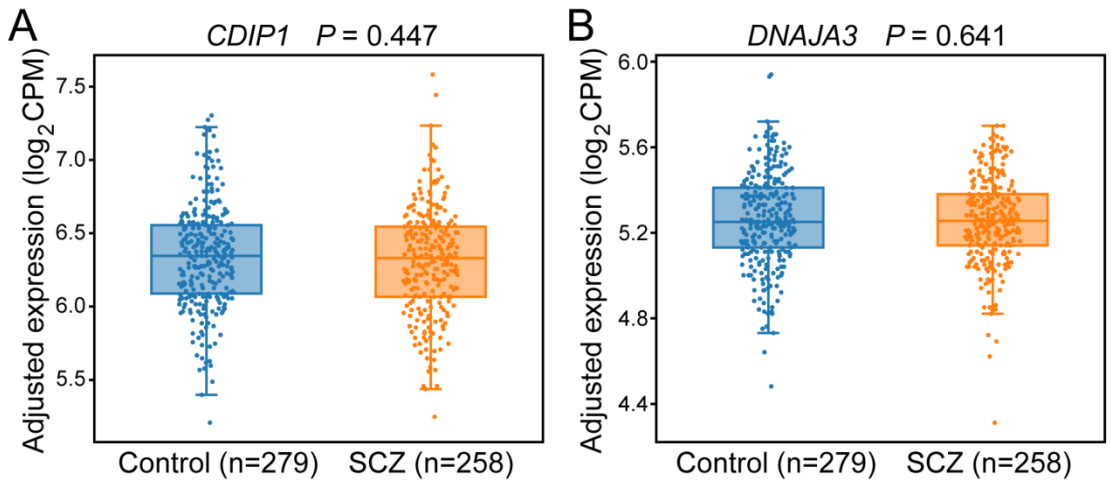
Supplementary Figure 11. Association significance between rs2270363 and its nearby genes in the human brain. (A) eQTL analysis of genes within 1 Mb window. Data were from LIBD, n = 412.[11] (**B**-**E**) The box-plot of eQTL analysis. A allele of rs2270363 corresponds to lower expression of *CORO7*, and corresponds to higher expression of *NMRAL1*, *DNAJA3*, *CDIP1*. Data were from LIBD, n = 404.[11] (**F**-**H**) qPCR validation of *DNAJA3*, *CDIP1* and *HMOX2* expression in rs2270363 knock-out HEK-293T cells. Unpaired two-tailed *Student's t*-test was used in this study. n = 3. Data are represented as mean ± SD. ns, not significant.



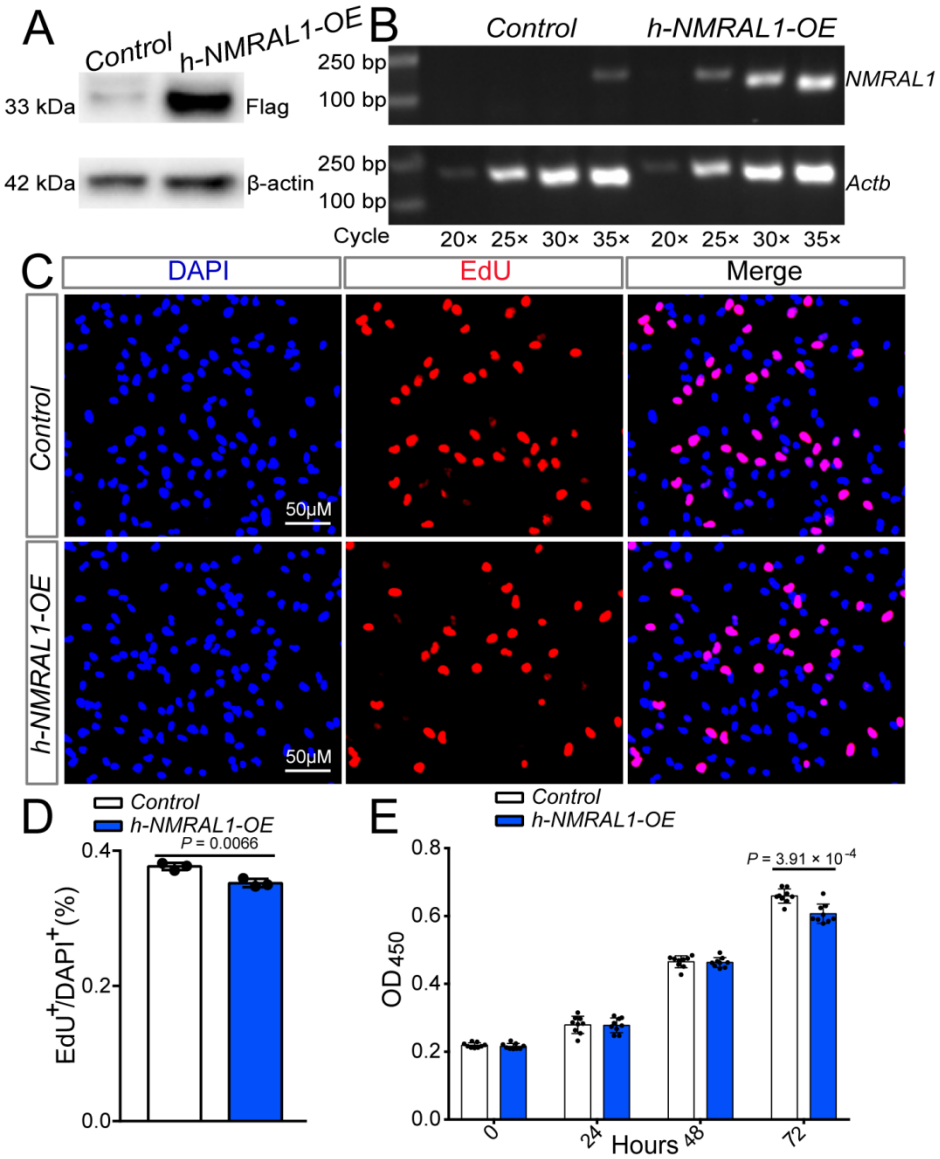
Supplementary Figure 12. *NMRAL1, CORO7, DNAJA3, CDIP1* and *HMOX2* are widely expressed in different tissues. (**A-C**) *NMRAL1*, *CORO7* and *DNAJA3* are widely expressed in various tissues, and the expression of *CDIP1* has low human brain regional specificity (**D**). (**E**) *HMOX2* is widely expressed in various tissues. Data were from GTEx Portal.[12]


A
NX
NMRAL1
B
NX
CORO7
C
NX
DNAJA3
D
NX
CDIP1
E
NX
HMOX2

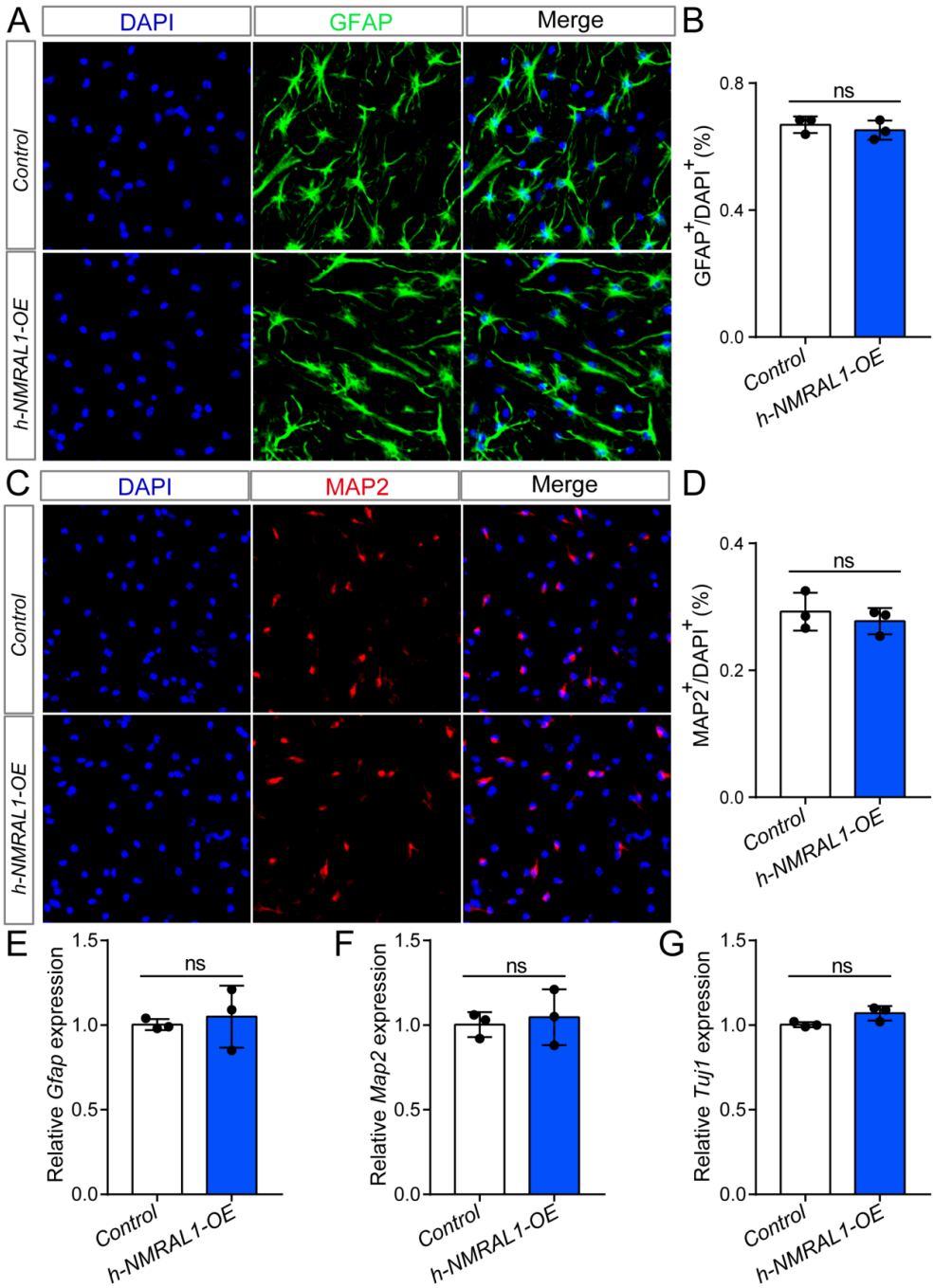
Supplementary Figure 13. *NMRAL1, CORO7, DNAJA3, CDIP1* and *HMOX2* are expressed in different cell types. (**A-E**) *NMRAL1*, *CORO7*, *DNAJA3*, *CDIP1* and *HMOX2* are widely expressed in various cell types, including SH-SY5Y and HEK293 cell lines. Data were from the Human Protein Atlas.[13]



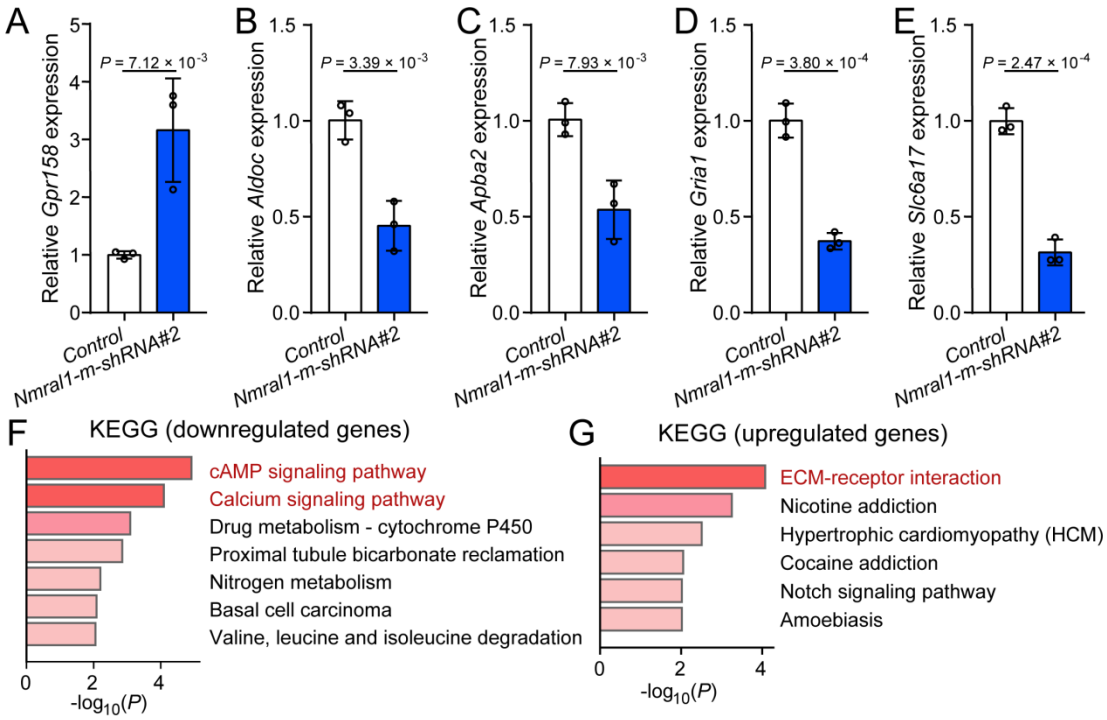
Supplementary Figure 14. The expression of *DNAJA3* and *CDIP1* did not show significant change in brains of SZ cases compared with controls. CMC expression data were extracted from SZDB2.0.[14]



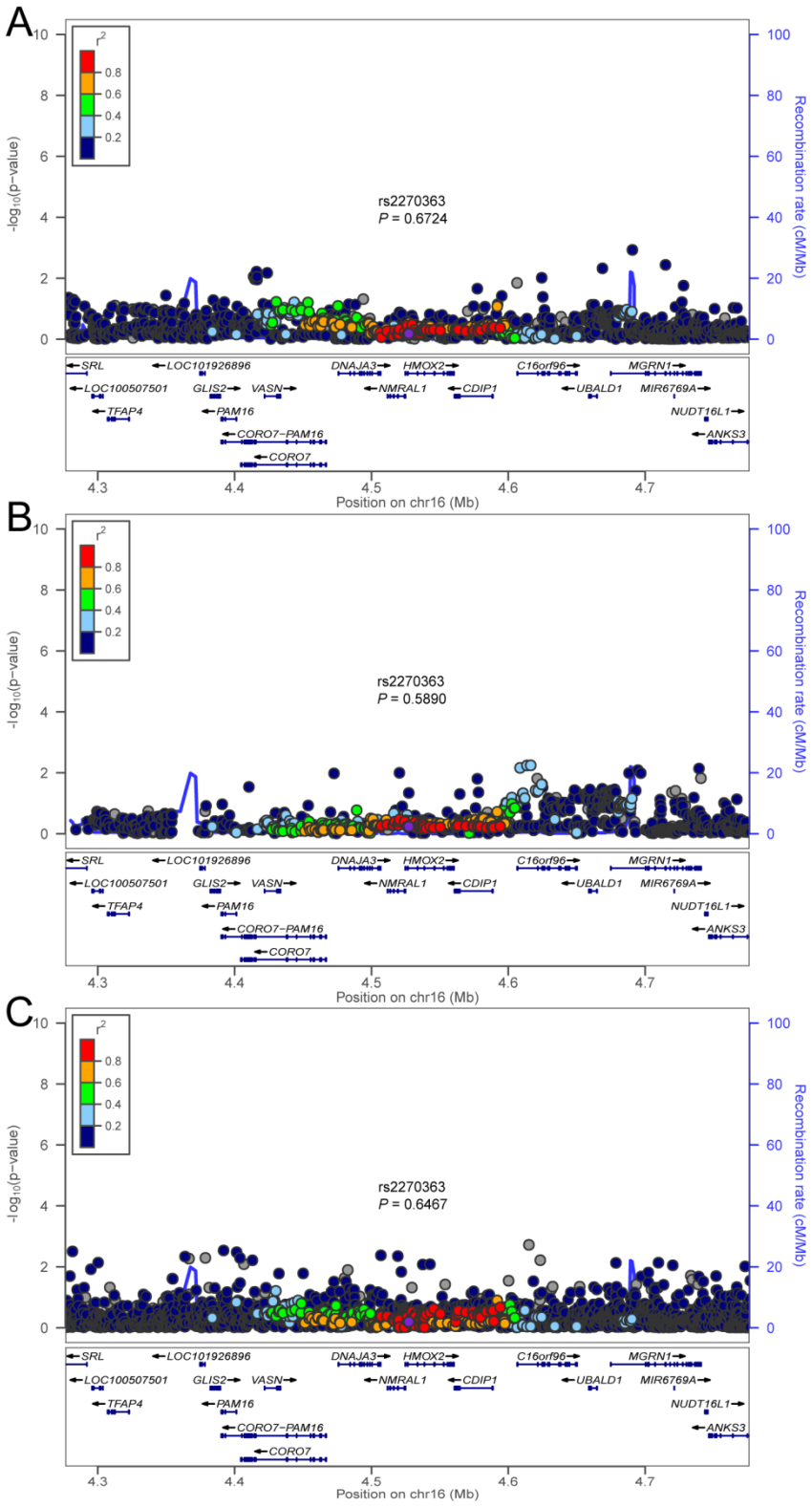
Supplementary Figure 15. *NMRAL1* overexpression significantly inhibited proliferation of the mNSCs. (A) Validation of *NMRAL1* overexpression in mouse NSCs with Western-blot and RT-PCR (**B**). (**C**) Immunofluorescence staining for EdU incorporation assay. Red indicates EdU positive cells, and DAPI was used to stain the nucleus (blue). (**D**) The quantification results of the EdU incorporation assay. (**E**) The results of CCK-8 assay. Data were measured at four different time points, 0, 24, 48 and 72 hours. Unpaired two-tailed *Student's t*-test was used to compare if the difference was significant. n = 3 for **B**-**D**; n = 9 for **E**. Data are represented as mean ± SD.



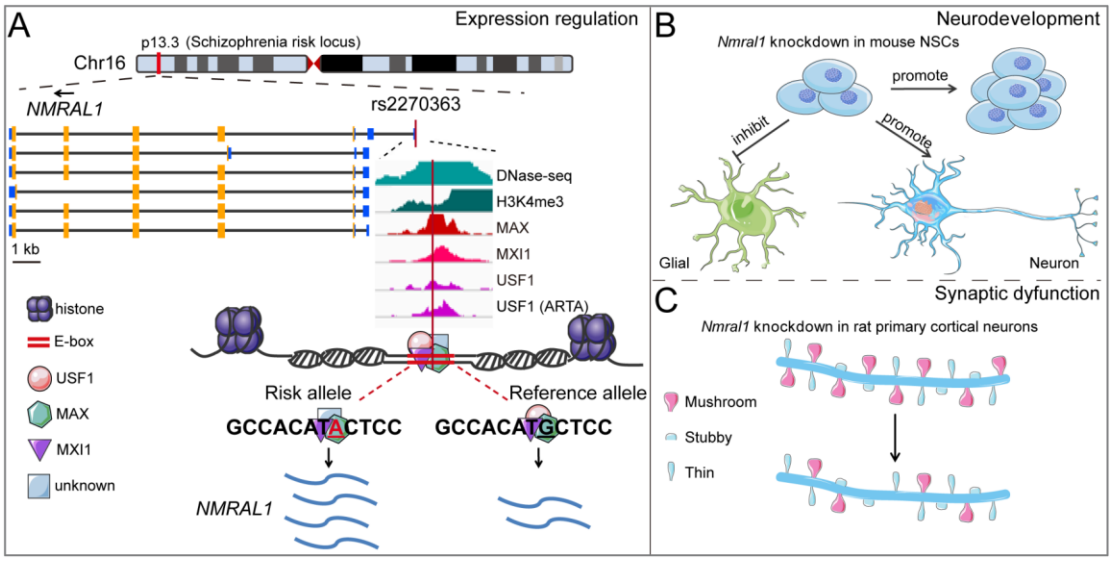
Supplementary Figure 16. *NMRAL1* overexpression did not affect differentiation of mNSCs. (**A**) Representative immunofluorescence staining images for glia cells differentiated from mNSCs. GFAP positive cells were marked with green and the nuclei were stained with DAPI. (**B**) Quantification for the proportion of GFAP positive glia cells. (**C**) Representative immunofluorescence staining images for neurons differentiated from mNSCs. MAP2 positive cells were marked with red and the nuclei were stained with DAPI. (**D**) Quantification for the proportion of MAP2 positive neurons. (**E-G**) qPCR showed the relative expression level of *Gfap*, *Map2* and *Tuj1* in control and *NMRAL1* overexpression mNSCs. Unpaired two-tailed *Student's t*-test was used in this study. n = 3. Data are represented as mean ± SD. ns, not significant.



Supplementary Figure 17. GO and KEGG analyses. (**A**-**E**) The results of qPCR verification. *P* values were calculated using two-tailed *Student's t*-test. n = 3 for **A**-**E**. (**F**, **G**) KEGG analyses of downregulated genes and upregulated genes.



Supplementary Figure 18. The Locus zoom plot showing the associations between variants in genomic region (500 kb) containing rs2270363 and ASD, ADHD and BPD. (**A**) The associations of genetic variants near rs2270363 (500 kb) and ASD.[15] (**B**) The associations of genetic variants near rs2270363 (500 kb) and ADHD.[16] (**C**) The associations of genetic variants near rs2270363 (500 kb) and BPD.[17]



Supplementary Figure 19. The working model of rs2270363 in schizophrenia pathogenesis. (**A**) The molecular regulatory mechanisms of rs2270363. rs2270363 lies in the E-box element of the promoter of *NMRAL1*. rs2270363 regulates *NMRAL1* expression through affecting TFs (USF1, MAX and MXI1) binding and promoter activity. The A allele is associated with higher expression of *NMRAL1*. (**B**, **C**) *Nmral1* affects neurodevelopment and synaptic morphogenesis. *Nmral1* knockdown affects proliferation and differentiation of mNSCs, and leads to decreased density of mushroom dendritic spines in rat primary cortical neurons.

Supplementary Table 1. The scoring scheme of the original RegulomeDB rank scores

Score	Supporting data
1a	eQTL + TF binding + matched TF motif + matched DNase Footprint + DNase peak
1b	eQTL + TF binding + any motif + DNase Footprint + DNase peak
1c	eQTL + TF binding + matched TF motif + DNase peak
1d	eQTL + TF binding + any motif + DNase peak
1e	eQTL + TF binding + matched TF motif
1f	eQTL + TF binding / DNase peak
2a	TF binding + matched TF motif + matched DNase Footprint + DNase peak
2b	TF binding + any motif + DNase Footprint + DNase peak
2c	TF binding + matched TF motif + DNase peak
3a	TF binding + any motif + DNase peak
3b	TF binding + matched TF motif
4	TF binding + DNase peak
5	TF binding or DNase peak
6	Motif hit
7	Other

Supplementary Table 2. The probes used for EMSA

Probes	Sequence (5' to 3')
rs2270363-EMSA-AF	AGACCTGGGGAAAGCCACAT**A**CTCCGCTTCCCGCC
rs2270363-EMSA-AR	GGCGGGAAGCGGAG**T**ATGTGGCTTTCCCCAGGTCT
rs2270363-EMSA-GF	AGACCTGGGGAAAGCCACAT**G**CTCCGCTTCCCGCC
rs2270363-EMSA-GR	GGCGGGAAGCGGAG**C**ATGTGGCTTTCCCCAGGTCT

The position of rs2270363 was marked with red.

Supplementary Table 3. shRNAs used for gene knockdown

sgRNAs and shRNAs	Sequence (5' to 3')
*sg*RNA-F	caccGGGGCGGGAAGCGGAGTATG
*sg*RNA-R	aaacCATACTCCGCTTCCCGCCCC
USF1-shRNA#1-F	CCGGGCCGAGACAAGATCAACAACTCTCGAGAGTTGTTGATCTTGTCTCGGCTTTTTG
USF1-shRNA#1-R	AATTCAAAAAGCCGAGACAAGATCAACAACTCTCGAGAGTTGTTGATCTTGTCTCGGC
USF1-shRNA#2-F	CCGGCCACGGATTAGAGGTCGTCATCTCGAGATGACGACCTCTAATCCGTGGTTTTTG
USF1-shRNA#2-R	AATTCAAAAACCACGGATTAGAGGTCGTCATCTCGAGATGACGACCTCTAATCCGTGG
MAX-shRNA#1-F	CCGGACACACACCAGCAAGATATTGCTCGAGCAATATCTTGCTGGTGTGTGTTTTTTG
MAX-shRNA#1-R	AATTCAAAAAACACACACCAGCAAGATATTGCTCGAGCAATATCTTGCTGGTGTGTGT
Nmral1-m[a]*-shRNA#1-F*	CCGGCATTCAGGATTCGAGTGGTAACTCGAGTTACCACTCGAATCCTGAATGTTTTTG
Nmral1-m-shRNA#1-R	AATTCAAAAACATTCAGGATTCGAGTGGTAACTCGAGTTACCACTCGAATCCTGAATG
Nmral1-m-shRNA#2-F	CCGGGCCTTGCTATTTCGAGAATCTCTCGAGAGATTCTCGAAATAGCAAGGCTTTTTG
Nmral1-m-shRNA#2-R	AATTCAAAAAGCCTTGCTATTTCGAGAATCTCTCGAGAGATTCTCGAAATAGCAAGGC
Nmral1-rat-shRNA1-F	TGCCTCCATTATGTAGTGTATTCAAGAGATACACTACATAATGGAGGCTTTTTTC
Nmral1-rat-shRNA1-R	TCGAGAAAAAAGCCTCCATTATGTAGTGTATCTCTTGAATACACTACATAATGGAGGCA
Nmral1-rat-shRNA2-F	TGCCTTGCTATTTCGAGAACTTCAAGAGAGTTCTCGAAATAGCAAGGCTTTTTTC
Nmral1-rat-shRNA2-R	TCGAGAAAAAAGCCTTGCTATTTCGAGAACTCTCTTGAAGTTCTCGAAATAGCAAGGCA

[a]This shRNA targeting *Nmral1* for mouse.

Supplementary Table 4. Primers used for *NMRAL1* overexpression

Primers	Sequence (5' to 3')
h-NMRAL1-OE-F	ctcatcattttggcaaagaattcATGGTGGACAAGAAACTGG
h-NMRAL1-OE-R	ctctagagtcgcggccgcTCACTTATCGTCGTCATCCTTGTAATCCAGCAGGTTGAAGTCCC

Supplementary Table 5. Primers used for qPCR

Primers	Sequence (5’ to 3’)
USF1-QPCR-F	CTGCTGTTGTTACTACCCAGG
USF1-QPCR-R	TCTGACTTCGGGGAATAAGGG
MAX-QPCR-F	GAGGTGGAGAGCGACGAAGA
MAX-QPCR-R	CTGTCTTTGATGTGGTCCCTACG
CORO7-human-QPCR-F	GCTGCCATTCAGACCTAGTCA
CORO7-human-QPCR-R	AGTCGCCAGAGTTTTACCGTC
HMOX2-human-QPCR-F	TCAGCGGAAGTGGAAACCTC
HMOX2-human-QPCR-R	AGAAGTCCTTGACAAACTGGGT
DNAJA3-human-QPCR-F	TTTGGCGAGTTCTCATCCTCT
DNAJA3-human-QPCR-R	TTGCAGCTTGATTGAATGTCAAC
CDIP1-human-QPCR-F	ATGTGATCTGGGCTGCTGC
CDIP1-human-QPCR-R	GCGCTTGTACGTGTAGATGTAGG
NMRAL1-human-QPCR-F	AAATACGTCGGCCAGAACATC
NMRAL1-human-QPCR-R	AGCTTTTCGTAGTCCTCAGGAG
Nmral1-mouse-QPCR-F	GGCCCGTGCATTGCTAGAA
Nmral1-mouse-QPCR-R	CGTCTCCCAGTAATTGGTCAC
Nmral1-rat-QPCR-F	CGACTGCCTTGCTATTTCGA
Nmral1-rat-QPCR-R	GTCACTCACAGACATTCCATCCA
ACTB-qPCR-F	CATGTACGTTGCTATCCAGGC
ACTB-qPCR-R	CTCCTTAATGTCACGCACGAT
Actb-qPCR-F	GGCTGTATTCCCCTCCATCG
Actb-qPCR-R	CCAGTTGGTAACAATGCCATGT
Gfap-qPCR-F	CGGAGACGCATCACCTCTG
Gfap-qPCR-R	AGGGAGTGGAGGAGTCATTCG
Map2-qPCR-F	GCCAGCCTCAGAACAAACA
Map2-qPCR-R	GCTCAGCGAATGAGGAAGGA
Tuj1-qPCR-F	TAGACCCCAGCGGCAACTAT
Tuj1-qPCR-R	GTTCCAGGTTCCAAGTCCACC
Gpr158-qPCR-F	GGAAAGTCACCGTCCCTA
Gpr158-qPCR-R	GAAGTTGGTGGCGTGT
Apba2-qPCR-F	TCACCTACTACATCCGCTAC
Apba2-qPCR-R	CTCCTGACACTCATCGGT
Aldoc-qPCR-F	GTCAGAGTGAAGAGGAGGCT
Aldoc-qPCR-R	CAGCATTATCCCTTTGTCCT
Slc6a17-qPCR-F	CGGAAACTACTTTGTCACCAT
Slc6a17-qPCR-R	GATGTTCTCAAGGATGACGAT
Gria1-qPCR-F	ACATCGTCACGACTATCCTCGA
Gria1-qPCR-R	AGTAGCCCTCATAGCGGTCATT

Supplementary Table 6. SNPs locations and RegulomeDB annotation

CHR	BP	SNP ID	the RegulomeDB score
chr16	**4526292**	**rs2270363**	**1f**
chr22	**42486723**	**rs1801311**	**1f**
chr22	**42475703**	**rs2269524**	**1f**
chr4	**103749516**	**rs223387**	**1f**
chr5	**140187322**	**rs3822346**	**1f**
chr2	**73461553**	**rs41285977**	**2a**
chr2	**110371692**	**rs7579996**	**2a**
chr22	**50311973**	**rs9616378**	**2a**
chr15	**85114447**	**rs12912934**	**2b**
chr15	**89878391**	**rs2856268**	**2b**
chr1	**243419429**	**rs3904682**	**2b**
chr8	**144815612**	**rs7012106**	**2b**
chr22	**42475568**	**rs8135801**	**2b**
chr14	**104314183**	**rs10083367**	**3a**
chr14	**104314182**	**rs10083370**	**3a**
chr1	**8484228**	**rs159961**	**4**
chr4	**103748797**	**rs223390**	**4**
chr19	**50169020**	**rs2304204**	**4**
chr3	**50340996**	**rs28365992**	**4**
chr12	**123849921**	**rs28594416**	**4**
chr3	**180630191**	**rs34455584**	**4**
chr4	**170540857**	**rs3797040**	**4**
chr11	**82996986**	**rs494791**	**4**
chr12	**123849272**	**rs60754073**	**4**
chr15	**40583560**	**rs62021888**	**4**
chr12	**123752637**	**rs76514049**	**4**
chr5	**140027216**	**rs778593**	**4**
chr2	**200716119**	**rs796364**	**4**
chr6	**87861585**	**rs9362397**	**4**
chr22	**50312979**	**rs9616382**	**4**
chr1	150039678	rs2027349	1b
chr3	52833219	rs2535629	1b
chr8	8168413	rs7014953	1b
chr22	50297435	rs1321	1f
chr19	19611550	rs3794993	1f
chr16	30017895	rs3814880	1f
chr16	4506232	rs4786494	1f
chr15	85177297	rs10795	2a
chr1	205110946	rs12136320	2a
chr17	17997547	rs2974999	2a
chr4	103605916	rs393223	2a

CHR	BP	SNP ID	the RegulomeDB score
chr12	123722833	rs78866909	2a
chr10	104387735	rs10786700	2b
chr16	68418792	rs11861362	2b
chr8	10009949	rs11986122	2b
chr8	10009952	rs11993089	2b
chr3	180630565	rs1805579	2b
chr19	50168871	rs2304206	2b
chr2	233792109	rs2675960	2b
chr1	8468278	rs301792	2b
chr1	214163069	rs340836	2b
chr16	89699664	rs4785581	2b
chr22	50317187	rs7410601	2b
chr22	42567090	rs9306356	2b
chr6	143646783	rs9373388	2b
chr8	89566903	rs1352318	2c
chr22	50178917	rs138833	2c
chr10	104928914	rs12416331	3a
chr4	103974122	rs13113099	3a
chr2	200787719	rs281759	3a
chr15	78894759	rs3743078	3a
chr1	150130096	rs78681982	3a
chr5	101769984	rs10038801	4
chr17	2143460	rs10852932	4
chr11	83048096	rs11233566	4
chr6	96464060	rs117178087	4
chr1	205035455	rs16937	4
chr2	220049128	rs2385395	4
chr2	208481637	rs2551945	4
chr7	24595020	rs2711116	4
chr22	42527471	rs28633410	4
chr6	28779878	rs3131340	4
chr3	53270344	rs3773744	4
chr3	53270424	rs3773745	4
chr6	143658860	rs3804536	4
chr6	108944165	rs3813498	4
chr12	123743436	rs4759413	4
chr8	8639740	rs486781	4
chr17	17897739	rs4924832	4
chr15	89902032	rs4932217	4
chr22	42380001	rs5751195	4
chr22	42511002	rs6002621	4
chr17	20170608	rs61660810	4
chr3	53376669	rs6795127	4

CHR	BP	SNP ID	the RegulomeDB score
chr5	153688215	rs6871683	4
chr8	8168222	rs6992091	4
chr1	150135291	rs72694957	4
chr15	84915246	rs72748702	4
chr12	110662327	rs7304243	4
chr17	19912679	rs78532287	4
chr5	127149861	rs790475	4
chr5	137775581	rs982085	4
chr5	49636753	rs10940235	5
chr5	153498135	rs58848914	5
chr8	89516441	rs716881	5
chr22	50278642	rs910800	5
chr8	9573154	rs12114661	7

The bold font indicates the 30 SNPs located in promoter regions and marked with H3K4me3 in UCSC genome browser (http://genome.ucsc.edu/). The red font indicate the 5 prioritized high-confidence (RegulomeDB rank score < 2) functional SNPs in the promoter region. Data were from RegulomeDB (https://www.regulomedb.org/).[7]

Supplementary Table 7. The eQTL genes of the 5 prioritized SNPs

SNP	Gene Symbol	*P*-value (CMC)	FDR (≤0.05, CMC)	*P*-value (LIBD)	FDR (≤0.05, LIBD)
rs2270363	*CDIP1*	3.05E-07	4.40E-05	1.43E-05	1.11E-03
	CORO7	2.06E-10	5.06E-08	1.72E-17	7.46E-15
	DNAJA3	8.22E-31	1.17E-27	2.21E-06	2.09E-04
	NMRAL1	1.62E-10	4.05E-08	4.88E-16	1.87E-13
rs2269524	*CYP2D7P*	NA	NA	8.48E-09	1.29E-06
	FAM109B	1.02E-06	9.73E-05	7.35E-16	2.77E-13
	LINC00634	2.33E-08	3.01E-06	1.01E-09	1.78E-07
	NAGA	5.33E-18	2.03E-15	1.04E-26	8.81E-24
	RP1-257I20.14	NA	NA	3.98E-05	2.71E-03
	RP4-669P10.16	NA	NA	2.93E-22	1.86E-19
	SMDT1	NA	NA	3.91E-06	3.49E-04
	WBP2NL	3.63E-06	3.09E-04	NA	NA
rs1801311	*CYP2D7P*	NA	NA	4.17E-12	1.03E-09
	FAM109B	1.46E-06	1.34E-04	6.23E-17	2.58E-14
	LINC00634	2.40E-08	3.08E-06	1.90E-08	2.73E-06
	NAGA	4.06E-17	1.43E-14	4.45E-29	4.42E-26
	RP1-257I20.14	NA	NA	7.28E-06	6.07E-04
	RP4-669P10.16	NA	NA	3.51E-23	2.37E-20
	WBP2NL	6.62E-06	5.33E-04	NA	NA
rs223387	*BDH2*	5.69E-07	3.98E-05	NA	NA
	CENPE	1.21E-03	3.76E-02	NA	NA
	CISD2	6.80E-04	2.35E-02	NA	NA
	KRT8P46	2.32E-14	4.62E-12	NA	NA
	LRRC37A15P	2.84E-13	4.96E-11	NA	NA
	SLC9B1	2.29E-35	1.51E-32	NA	NA
rs3822346	*ANKHD1-EIF4EBP3*	2.02E-04	9.40E-03	NA	NA
	CD14	NA	NA	1.53E-04	8.55E-03
	IK	8.96E-04	3.20E-02	3.49E-05	2.41E-03
	NDUFA2	3.14E-05	1.88E-03	NA	NA
	PCDHA1	NA	NA	2.62E-07	3.03E-05
	PCDHA10	1.96E-14	4.50E-12	1.30E-13	3.86E-11
	PCDHA13	3.90E-19	1.49E-16	3.21E-14	1.02E-11
	PCDHA2	3.69E-08	3.96E-06	3.95E-06	3.53E-04
	PCDHA7	9.78E-20	4.04E-17	1.30E-14	4.29E-12
	PCDHA8	2.43E-18	8.29E-16	5.03E-32	5.90E-29
	PCDHA9	NA	NA	6.19E-18	2.80E-15
	PCDHAC2	5.85E-04	2.26E-02	NA	NA
	PCDHB3	1.69E-04	8.13E-03	NA	NA
	PCDHB9	8.28E-04	3.00E-02	NA	NA
	SRA1	2.53E-07	2.34E-05	NA	NA

SNP	Gene Symbol	*P*-value (CMC)	FDR (≤0.05, CMC)	*P*-value (LIBD)	FDR (≤0.05, LIBD)
	TMCO6	NA	NA	8.82E-05	5.37E-03
rs3822346	*WDR55*	9.48E-04	3.35E-02	NA	NA
	ZMAT2	1.30E-10	2.00E-08	1.10E-08	1.65E-06

Data were from CMC (N = 467) and LIBD (N = 412) brain eQTL datasets.[8, 11]

Supplementary Table 8. Differential expression analysis of the eQTL genes of the 5 selected functional SNPs

SNP	Gene_symbol	logFC	P-value	FDR
rs2270363	*CDIP1*	-0.026	4.47E-01	6.92E-01
	CORO7	**-0.086**	**2.46E-02**	**1.65E-01**
	DNAJA3	-0.010	6.41E-01	8.25E-01
	NMRAL1	**-0.104**	**3.98E-03**	**7.18E-02**
rs2269524	*CYP2D7P*	NA	NA	NA
	FAM109B	**-0.108**	**2.20E-02**	**1.57E-01**
	LINC00634	**-0.172**	**3.43E-04**	**2.43E-02**
	NAGA	0.015	6.76E-01	8.44E-01
	RP1-257I20.14	NA	NA	NA
	RP4-669P10.16	NA	NA	NA
	SMDT1	NA	NA	NA
	WBP2NL	0.036	4.21E-01	6.73E-01
rs1801311	*CYP2D7P*	NA	NA	NA
	FAM109B	**-0.108**	**2.20E-02**	**1.57E-01**
	LINC00634	**-0.172**	**3.43E-04**	**2.43E-02**
	NAGA	0.015	6.76E-01	8.44E-01
	RP1-257I20.14	NA	NA	NA
	RP4-669P10.16	NA	NA	NA
	WBP2NL	0.036	4.21E-01	6.73E-01
rs223387	*BDH2*	-0.033	3.63E-01	6.24E-01
	CENPE	0.069	9.58E-02	3.22E-01
	CISD2	-0.006	8.44E-01	9.28E-01
	KRT8P46	0.027	4.89E-01	7.23E-01
	LRRC37A15P	0.013	7.38E-01	8.73E-01
	SLC9B1	-0.041	3.60E-01	6.21E-01
rs3822346	*ANKHD1-EIF4EBP3*	0.077	5.23E-02	2.41E-01
	CD14	0.031	7.63E-01	8.86E-01
	IK	-0.014	6.79E-01	8.45E-01
	NDUFA2	**-0.124**	**1.48E-02**	**1.32E-01**
	PCDHA1	-0.020	6.39E-01	8.24E-01
	PCDHA10	-0.034	6.98E-01	8.55E-01
	PCDHA13	-0.074	8.43E-02	3.04E-01
	PCDHA2	0.039	3.16E-01	5.84E-01
	PCDHA7	0.093	7.48E-02	2.88E-01
	PCDHA8	0.006	9.03E-01	9.57E-01
	PCDHA9	-0.003	9.69E-01	9.86E-01
	PCDHAC2	0.026	4.05E-01	6.61E-01
	PCDHB3	0.009	8.00E-01	9.05E-01
	PCDHB9	0.017	6.54E-01	8.32E-01
	SRA1	**-0.059**	**1.85E-02**	**1.46E-01**

SNP	Gene_symbol	logFC	P-value	FDR
	TMCO6	-0.026	3.69E-01	6.29E-01
rs3822346	*WDR55*	-0.016	4.48E-01	6.93E-01
	ZMAT2	-0.069	1.08E-01	3.40E-01

[a]The eQTL genes of 5 SNPs were obtained from CMC and LIBD, (FDR ≤ 0.05).

The bold font represents the genes down-regulated in SZ cases compared with controls.

Supplementary Table 9. rs2269524 is in high linkage disequilibrium with rs1801311 in different populations

Population	Focus Variant	Variant 2	r^2	D'
Bengali in Bangladesh	rs2269524	rs1801311	0.97	1
Chinese Dai in Xishuangbanna, China	rs2269524	rs1801311	0.95	1
Utah residents with Northern and Western European ancestry	rs2269524	rs1801311	0.98	1
Han Chinese in Bejing, China	rs2269524	rs1801311	0.97	1
Southern Han Chinese, China	rs2269524	rs1801311	1.00	1
Finnish in Finland	rs2269524	rs1801311	1.00	1
British in England and Scotland	rs2269524	rs1801311	1.00	1
Gujarati Indian in Houston, TX	rs2269524	rs1801311	1.00	1
Iberian populations in Spain	rs2269524	rs1801311	0.98	1
Indian Telugu in the UK	rs2269524	rs1801311	0.98	1
Japanese in Tokyo, Japan	rs2269524	rs1801311	1.00	1
Kinh in Ho Chi Minh City, Vietnam	rs2269524	rs1801311	0.96	1
Mexican Ancestry in Los Angeles, California	rs2269524	rs1801311	1.00	1
Peruvian in Lima, Peru	rs2269524	rs1801311	1.00	1
Punjabi in Lahore, Pakistan	rs2269524	rs1801311	1.00	1
Sri Lankan Tamil in the UK	rs2269524	rs1801311	0.98	1
Toscani in Italy	rs2269524	rs1801311	0.98	1

r^2 indicates the linkage disequilibrium value between these two SNPs. Data were from Ensembl.

Supplementary Table 10. Genotypic frequency of rs2270363 in tissue samples (258 cases, 279 controls) used for differential expression analysis (CMC).

Genotype	Samples		Genotypic frequency	
	SZ cases	Controls	SZ cases	Controls
AA	15	24	0.0586	0.0860
AG	103	91	0.4023	0.3262
GG	138	164	0.5391	0.5878

Data were from CMC.[8] Two cases have no genotype data.

Supplementary Table 11. Primers used for SNaPShot genotyping

Primers	Sequence (5' to 3')
rs2270363-F-252	AGTGTTCCTGCCGAAGCAC
rs2270363-R-252	CAGCCTCCAGTCAGCGTG
Ex-rs2270363 [a]	TTTTTTTTTT AGACCTGGGG AAAGCCACAT

[a]The extension primers were used for genotype.

Supplementary Table 12. Primers used for dual-luciferase reporter gene assays

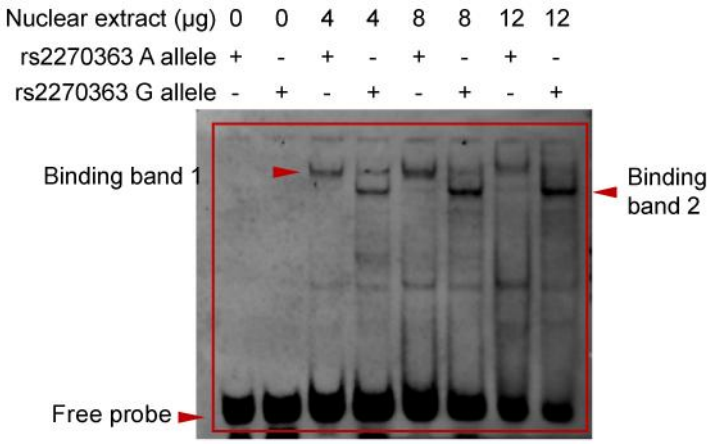
Primers	Sequence (5' to 3')
rs2270363-F1-861-wt[a]	CCAGAACATTTCTCTATCGATAGGTACCCCTATAGAGATAAGGGCATTAGAGG
rs2270363-R1-861-wt[a]	ACTTAGATCGCAGATCTCGAGGTGGATGTACAAGGATTAAACCAC
rs2270363-F2-861-wt	GGCCTAACTGGCCGGTACCGTGGATGTACAAGGATTAAACCAC
rs2270363-R2-861-wt	ACAGTACCGGATTGCCAAGCTTCCTATAGAGATAAGGGCATTAGAGG
rs2270363-F-861-mut[b]	GCCACAT G CTCCGCTTCCCGCCCCC
rs2270363-R-861-mut[b]	AGCGGAG C ATGTGGCTTTCCCCAGGTCT

[a]This pair of primers were used to amplify the DNA fragments containing rs2270363.
[b]This pair of primers were used for PCR-mediated site mutation.

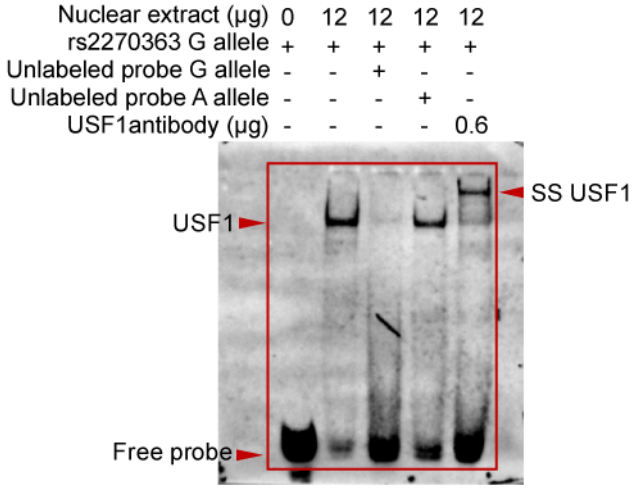
References

1. Li YF, Ma CG, Li WQ, Yang YF, Li XY, Liu JW *et al.* A missense variant in NDUFA6 confers schizophrenia risk by affecting YY1 binding and NAGA expression. *Mol Psychiatr* 2021**:** DOI: 10.1038/s41380-41021-01125-x.
2. Wang JY, Li XY, Li HJ, Liu JW, Yao YG, Li M *et al.* Integrative analyses followed by functional characterization reveal TMEM180 as a schizophrenia risk gene. *Schizophrenia bulletin* 2021**:** DOI: 10.1093/schbul/sbab1032.
3. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* 2014; **15**(12)**:** 550.
4. Zhou YY, Zhou B, Pache L, Chang M, Khodabakhshi AH, Tanaseichuk O *et al.* Metascape provides a biologist-oriented resource for the analysis of systems-level datasets. *Nat Commun* 2019; **10:** 1523.
5. Lam M, Chen CY, Li ZQ, Martin AR, Bryois J, Ma XX *et al.* Comparative genetic architectures of schizophrenia in East Asian and European populations. *Nat Genet* 2019; **51**(12)**:** 1670-1678.
6. Huo YX, Li SW, Liu JW, Li XY, Luo XJ. Functional genomics reveal gene regulatory mechanisms underlying schizophrenia risk. *Nat Commun* 2019; **10:** 670.
7. Boyle AP, Hong EL, Hariharan M, Cheng Y, Schaub MA, Kasowski M *et al.* Annotation of functional variation in personal genomes using RegulomeDB. *Genome Res* 2012; **22**(9)**:** 1790-1797.
8. Fromer M, Roussos P, Sieberts SK, Johnson JS, Kavanagh DH, Perumal TM *et al.* Gene expression elucidates functional impact of polygenic risk for schizophrenia. *Nat Neurosci* 2016; **19**(11)**:** 1442-1453.
9. Nott A, Holtman IR, Coufal NG, Schlachetzki JCM, Yu M, Hu R *et al.* Brain cell type-specific enhancer-promoter interactome maps and disease-risk association. *Science* 2019; **366**(6469)**:** 1134-1139.
10. Davis CA, Hitz BC, Sloan CA, Chan ET, Davidson JM, Gabdank I *et al.* The Encyclopedia of DNA elements (ENCODE): data portal update. *Nucleic Acids Res* 2018; **46**(D1)**:** D794-D801.
11. Jaffe AE, Straub RE, Shin JH, Tao R, Gao Y, Collado-Torres L *et al.* Developmental and genetic regulation of the human cortex transcriptome illuminate schizophrenia pathogenesis. *Nat Neurosci* 2018; **21**(8)**:** 1117-1125.
12. Ardlie KG, DeLuca DS, Segre AV, Sullivan TJ, Young TR, Gelfand ET *et al.* The Genotype-Tissue Expression (GTEx) pilot analysis: Multitissue gene regulation in humans. *Science* 2015; **348**(6235)**:** 648-660.
13. Persson A, Hober S, Uhlen M. A human protein atlas based on antibody proteomics. *Curr Opin Mol Ther* 2006; **8**(3)**:** 185-190.
14. Wu Y, Li XY, Liu JW, Luo XJ, Yao YG. SZDB2.0: an updated comprehensive resource for schizophrenia research. *Hum Genet* 2020; **139**(10)**:** 1285-1297.
15. Grove J, Ripke S, Als TD, Mattheisen M, Walters RK, Won H *et al.* Identification of common genetic risk variants for autism spectrum disorder. *Nat Genet* 2019; **51**(3)**:** 431-444.
16. Demontis D, Walters RK, Martin J, Mattheisen M, Als TD, Agerbo E *et al.* Discovery of the first genome-wide significant risk loci for attention deficit/hyperactivity disorder. *Nat Genet* 2019; **51**(1)**:** 63-75.
17. Stahl EA, Breen G, Forstner AJ, McQuillin A, Ripke S, Trubetskoy V *et al.* Genome-wide association study identifies 30 loci associated with bipolar disorder. *Nat Genet* 2019; **51**(5)**:** 793-803.

The unmodified EMSA blots



Original EMSA blots of Fig. 1I

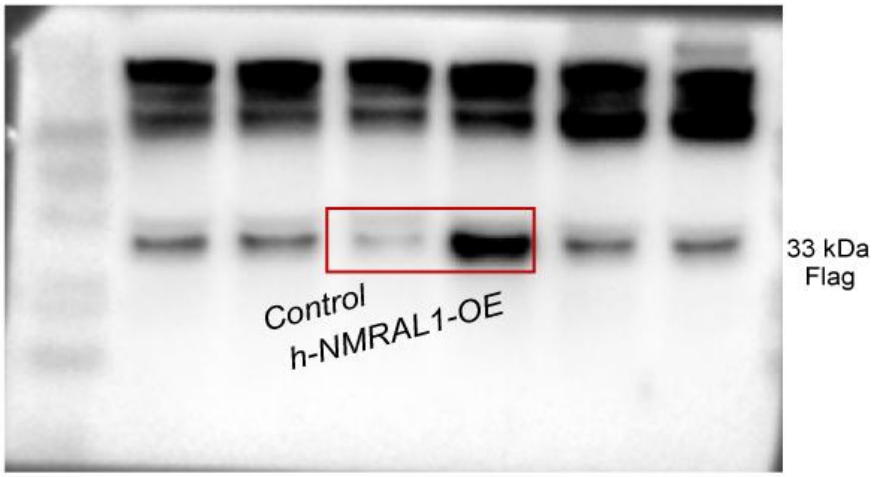


Original EMSA blots of Fig. 1J

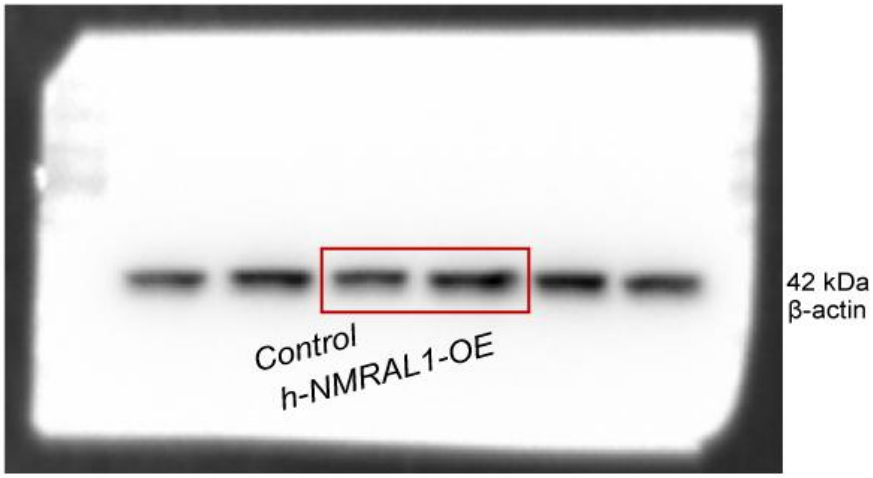
The unmodified blots of Western-blot

NMRAL1-1×FLAG and β-actin protein expression in mouse NSCs

NMRAL1-1×FLAG



β-actin

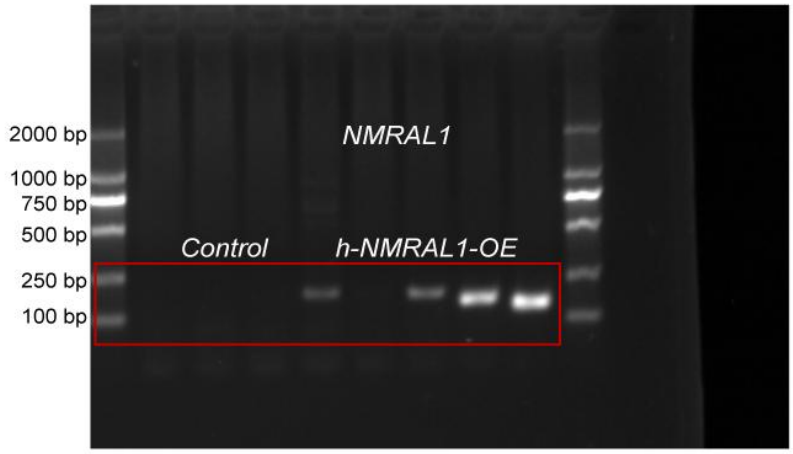


Original blots of Supplementary Fig. 15A

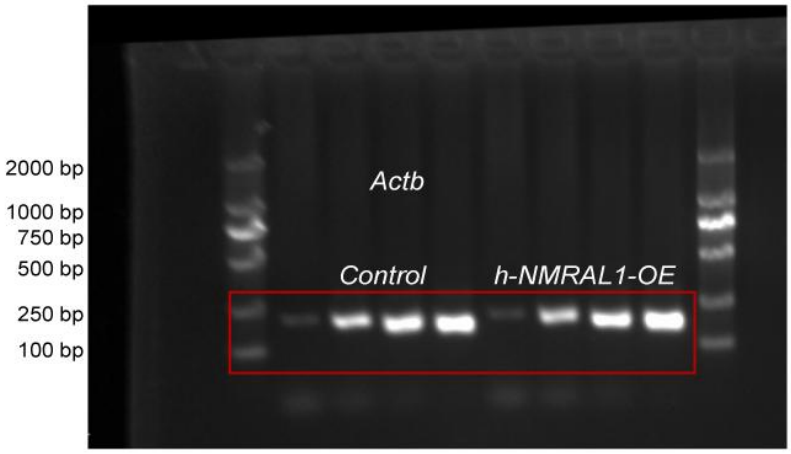
The unmodified gels of RT-PCR

NMRAL1 and *Actb* mRNA expression in mouse NSCs

NMRAL1



Actb



Original gels of Supplementary Fig. 15B